Deep Learning Algorithms to Isolate and Quantify the Structures of the Anterior Segment in Optical Coherence Tomography Images

Tan Hung Pham[1,3], Sripad Krishna Devalla[1], Aloysius Ang[2], Soh Zhi Da[3], Alexandre Thiery[4], Craig Boote[1,5,6], Ching-Yu Cheng[2,3,7*], Michael J.A Girard[1,3*], Victor Koh[8*]

1. Department of Biomedical Engineering, Faculty of Engineering, National University of Singapore, Singapore.
2. Yong Loo Lin School of Medicine, National University of Singapore, Singapore.
3. Singapore Eye Research Institute, Singapore National Eye Centre, Singapore.
4. Department of Statistics and Applied Probability, National University of Singapore, Singapore
5. School of Optometry and Vision Sciences, Cardiff University, UK.
6. Newcastle Research & Innovation Institute, Singapore.
7. Ophthalmology & Visual Sciences Academic Clinical Program (Eye ACP), Duke-NUS Medical School, Singapore
8. Department of Ophthalmology, National University Hospital, Singapore.

***** All three are corresponding authors. All of them contributed equally.

Corresponding Author:

Victor Koh, MD, Ph.D.
Department of Ophthalmology
National University Hospital
1E Kent Ridge Rd, Singapore 119228
victor_koh@nuhs.edu.sg

Cheng Ching Yu, MD, Ph.D.
Singapore Eye Research Institute
20 College Road, The Academia, Singapore 169856
chingyu.cheng@duke-nus.edu.sg

Michaël J. A. Girard, Ph.D.
Ophthalmic Engineering & Innovation Laboratory
Department of Bioengineering
National University of Singapore
4 Engineering Drive 3, E4-04-08
Singapore 117583
mgirard@nus.edu.sg
+65 6516 5549
http://www.bioeng.nus.edu.sg/oeil/

Abstract word count: 240
Total word count: 3700
Total figure count: 5
Total table count: 3

Keywords: deep learning; glaucoma; anterior segment optical coherence tomography; scleral spur

Contributorship Statement: Tan Hung Pham designed the experiments, algorithms and wrote the paper. Sripad Krisna Devalla assisted in algorithm design. Aloysius Ang and Victor Koh assisted in data labelling and experiment design. Sripad Krisna Devalla, Aloysius Ang, Soh Zhi Da, Alexandre Thiery and Craig Boote edited the manuscript. Ching-Yu Cheng, Michael J.A Girard and Victor Koh supervised the study and edited the manuscript.

Synopsis

Deep neural networks enable fast and accurate automated isolation and quantification of important intraocular dimensions in anterior segment of the eye in optical coherence tomography images.

Abstract

Background/Aims:

Accurate isolation and quantification of intraocular dimensions in the anterior segment (AS) of the eye using optical coherence tomography (OCT) images is important in the diagnosis and treatment of many eye diseases, especially angle closure glaucoma.

Methods:

In this study, we developed a deep convolutional neural network (DCNN) for the localization of the scleral spur, moreover we introduced an information rich segmentation approach for this localization problem. An ensemble of DCNNs for the segmentation of anterior segment structures (iris, corneo-sclera shell, anterior chamber) was developed. Based on the results of two previous processes, an algorithm to automatically quantify clinically important measurements were created. 200 images from 58 patients (100 eyes) were used for testing.

Results:

With limited training data, the DCNN was able to detect the scleral spur on unseen ASOCT images as accurately as an experienced ophthalmologist on the given test dataset; and simultaneously isolated the anterior segment structures with a Dice coefficient of 95.7%. We then automatically extracted eight clinically relevant ASOCT measurements and proposed an automated quality check process that asserts the reliability of these measurements. When combined with an OCT machine capable of imaging multiple radial sections, the algorithms can provide a more complete objective assessment. The total segmentation and measurement time for a single scan is less than 2 seconds.

Conclusion:

This is an essential step toward providing a robust automated framework for reliable quantification of ASOCT scans, for applications in the diagnosis and management of angle closure glaucoma.

INTRODUCTION

Primary angle closure glaucoma (PACG) is a major type of glaucoma, in particular in Asia [1]. By 2020, the number of people affected by primary angle closure glaucoma (PACG) is estimated to be up to 23.4 million[1 2]. PACG is associated with a high rate of blindness [3 4] that is up to 5 times greater than primary open-angle glaucoma[5]. Therefore, an early diagnosis followed by effective management strategies is essential to reduce the damage to the optic nerve head tissues that could lead to irreversible vision loss [6]. Early diagnosis is crucial in the Asian population, given the higher prevalence of PACG compared to European and African populations [3 4 7].

The diagnosis of PACG is based on the status of the anterior chamber angle (ACA) [8-10]. While the gold standard for ACA assessment is dark-room indentation gonioscopy [11], the procedure requires direct contact with the eye and is highly dependent on the physician's expertise and the background illumination [11 12]. This can result in poor reproducibility and diagnostic accuracy. In contrast, anterior segment optical coherence tomography (ASOCT) imaging allows for an objective, fast and non-contact assessment of the ACA in a standardized dark-room environment [12 13]. However, current technology typically requires the manual identification and marking of the scleral spur location (SSL) (**Supplement Figure 1**) by a human grader before ACA measurements such as trabecular iris space area (TISA) and angle opening distance (AOD) can be measured to quantify the anterior chamber angle [14]. The introduction of this subjective human factor has been shown to introduce significant intra- and inter-observer variability [12-14]. The inconsistent labelling of SSL compromises the diagnosis and the monitoring of treatment effectiveness/disease severity in PACG [14]. In addition, with swept-source ASOCT imaging, there are up to 128 cross-sectional scans obtained per eye. To manually label each individual section in a timely manner would not be clinically viable, and therefore automated image processing algorithms are required.

Deep convolutional neural networks (DCNNs) have been shown to perform well with many medical imaging modalities [15-19], but their applications in ASOCT imaging are nascent. From the perspective of the current study, there are two relevant applications that can benefit from DCNNs, namely: object localization (for SSL detection) and segmentation (for classifying tissues such as the cornea and the iris). Traditional object detection and localization approaches in DCNNs are mainly based on classification and regression [20]. However, this approach requires a large number of labelled images to achieve robust automation [21]. Moreover, accurate landmark localization is critical for the diagnosis and management of PACG. Hence with limited training data, a traditional regression approach is not ideal in providing a high accuracy prediction. Frequently, in the medical context, it might not be feasible to obtain a large number of labelled images due to limited resources and time. This problem is exacerbated in certain ocular conditions that are relatively less common which may benefit from mass screening such as PACG. In addition, the reduced availability of

ASOCT images for eyes with PACG can be attributed to the lack of accessible equipment, cost, and clinical expertise.

In this study, we developed a custom hybrid DCNN inspired from widely used U-Net and full-resolution residual network (FRRnet) [22] for the localization of scleral spur, and the segmentation of the anterior segment structures (iris, corneo-sclera shell, anterior chamber). The hybrid DCNN leveraged the U-Net architecture to simultaneously exploit the local (i.e. tissue texture) and contextual (i.e. tissue spatial arrangement) information and exploited the FRRnet pathway to achieve precise localization. Further, we automatically extracted eight clinically relevant ASOCT measurements from the segmented structures. The aim of the work is to offer a robust and automated framework for the accurate localization of the scleral spur and quantification of the ASOCT structures for enhancing the diagnosis and management of PACG.

METHODS

ASOCT imaging

We included ASOCT images from patients examined at the Eye Surgery Centre, National University Hospital, Singapore. Prior informed consent was obtained for all patients. The study was conducted in accordance with the tenets of the World Medical Association's Declaration of Helsinki and had ethics approval from the National Healthcare Group Domain Specific Review Board (NHG 292015/00788). In total, ASOCT images from 100 patients (175 eyes) were included for analysis. The scans were obtained from the swept-source Casia SS-1000 ASOCT (Tomey Corporation, Nagoya, Japan). All the eyes in this study was part of a prospective cohort study which included only eyes with primary angle closure suspects and so all the eyes were phakic. For each eye, a 360-degree scan yielded up to 128 cross-sections of the anterior segment. We used 620 images from 42 patients (75 eyes) for training and another 200 images from 58 patients (100 eyes) for testing. Since each image contained two scleral spur instances, we further divided the images in half for scleral spur localization (**Supplementary Figure 2**). All the images used for testing were out-of-sample validation, meaning training and model tuning were being done entirely on training images. All results reported are from testing images.

Small landmark localization and ASOCT segmentation

The accurate localization of small landmark points using neural networks has always been challenging [23]. In the current study, we adopted a segmentation approach for both the landmark localization and the ASOCT segmentation. A MATLAB (R2018a, MathWorks Inc., Natick, MA) script was prepared to assist in labelling the SSL (landmark localization). Three definitions were used to locate the scleral spur: **1)** A change in curvature in the corneo-scleral interface; **2)** The posterior end of the trabecular meshwork; and **3)** The posterior end of a protruding structure along the cornea and sclera [14 24]. In each image, the following classes

were identified **(Supplementary Figure 2)**: focus region; attention region and the background. Out of the 620 training images, 420 were used for training and 200 were used for validation. The full 200 test images were used for testing.

FIJI[25] was used to obtain the manual segmentations of the ASOCT tissues. In each image, the following classes were identified **(Supplementary Figure 3)**: **(1)** the iris; **(2)** the corneo-sclera shell; **(3)** the anterior chamber; and the background. Due to limited human resource and the complicated procedure of tissue segmentation, we only had 126 training, 18 validation and 84 testing images.

The SSL labelling and the manual segmentations used for training the DCNNs were prepared by two trainers: a trained medical student (AA), and a trained observer (THP), both with more than two years of experience in ASOCT imaging.

The landmark localization and segmentation performance of the DCNNs on unseen ASOCT images were evaluated by three graders: the aforementioned trained observer (observer A; THP) and medical student (Observer B; AA), and a glaucoma fellowship trained ophthalmologist (Observer C; VK) with eight years of experience in in the management of PACG.

Quantification of ASOCT measurements

The ASOCT measurements could be automatically quantified once the scleral spur was defined and the anterior segment intraocular tissues segmented. The key structural measurements, including ACA, anterior chamber and iris-based measurements were automatically computed based on their definitions (**Table 1**).

Table 1. Definitions of important anterior segment optical coherence tomography measurements

Measurement	Definition
Anterior Chamber Depth (ACD)	Axial distance between corneal endothelium to anterior lens surface [26]
Lens Vault (LV)	Perpendicular distance from middle of the line connecting the scleral spurs to the anterior pole of the lens [27]
Anterior Chamber Width (ACW)	Distance between two scleral spurs [28]
Anterior Chamber Area (ACA)	Area bordered by posterior surface of the cornea, anterior surface of iris and anterior surface of the lens [29]
Angle Opening Distance (AOD)	Distance between the anterior iris surface and posterior corneal surface on a line perpendicular to the trabecular meshwork, a distance from the scleral spur (500µm, 750µm etc.) [30]

Trabecular Iris Space Area (TISA)	Area of a trapezoid created by the following boundaries: AOD of a distance from scleral spur (500µm, 750µm etc.), line from scleral spur perpendicular to plane of inner scleral wall to the iris, inner corneoscleral wall, iris surface [30]
Iris thickness (IT)	IT at a distance from the scleral spur or a relative distance in the iris (e.g.: middle of iris) [31]
Iris Curvature (ICurve)	Distance from iris greatest convexity point to the line between most central and most peripheral iris pigment epithelium [31]

Network training and architecture

In recent years, several research groups have successfully used U-Net and its variants [17 19 32 33] in medical image segmentation. The sequential downsampling and upsampling of images combined with skip connections [34] help in simultaneously extracting both the local (i.e., tissue texture) and contextual (i.e., tissue spatial arrangement) information. This allows U-Net style architectures to achieve very high levels of segmentation accuracy even when trained with limited training data [16 17 19]. Another promising but less explored DCNN in medical imaging applications is the FRRnet [22]. The network has two pathways: a full resolution path that helps in identifying precise boundaries and a multi-scale feature extraction pathway that is responsible for robust feature recognition. Also, the residual connections improve the gradient flow through the network [35]. By combining the information from both the pathways, the FRRnet was able achieve precise localization and robust feature recognition [22].

Many studies have demonstrated that an ensemble network that learned to combine the predictions of multiple DCNNs into a single predictive model offered a better accuracy than each of the networks separately [36 37]. When trained on the same training data as the individual DCNNs (weights of the individual DCNNs were frozen), the ensemble network learned to reduce the variance for each network, thus dramatically increasing the predictive power.

In this study, we developed FRRUnet (full resolution residual U-Net), a hybrid DCNN that exploited the inherent advantages of both the U-Net and the FRRnet. For the detection of the SSL, the FRRUnet was used, while an ensemble of the U-Net, FRRnet, and the FRRUnet was used for the segmentation of the ASOCT structures [**Supplementary Figure 4,5,6,7]**.

All three networks were trained end to end using an Adam optimizer [38] with a learning rate of 5e-5 without any scheduler, β1 of 0.9 and β2 of 0.999, and categorical cross entropy loss function [39]. All the convolution layers were activated with a leaky rectifier linear unit (ReLU) [40] activation function. A dropout layer with a probability of 0.5 was used after every building block to reduce the overfitting [41]. Given the limited size of the training dataset, the DCNNs' variance was increased through data augmentation techniques such as

rotation, width shift, height shift, shear, zoom, flip, brightness and contrast shift. The final U-Net, FRRnet, FRRUnet, and the ensemble network consisted of 7.80 million, 4.2 million, 4.2 million, and 1.7 thousand trainable parameters respectively. All networks were trained and tested on an NVIDIA GTX 1080 founder's edition GPU with CUDA v8.0 and cuDNN v5.1 acceleration. Using the given hardware configuration, for each ASOCT image the network was able to detect the SSL in 0.108 ± 0.0035 seconds and segment the ASOCT tissues in 0.324 ± 0.0018 seconds. The measurements were then automatically computed on a CPU (Intel Xeon at 2.1 GHz) in under 1.723 ± 0.287 seconds. It should be noted that measurement quantification can be accelerated by parallelism since each scan is independent.

Inter- and intra- observer tests

We performed an inter-observer agreement test to assess the reproducibility when identifying the scleral spur between three human observers: A – Trained non-expert, B – Trained medical student; C – Fellowship-trained glaucoma expert well-versed in ASOCT analysis and the software algorithm. The intra-observer agreement test assessed the extent of repeatability among the human observers and their comparison with the software algorithm. The time interval between image grading by the same observer was between 3 and 7 days. A paired t-test was used to measure the extent of agreement on-average and Bland-Altman plots were used to depict the limit of agreement (±1.96 SD) and the distribution of discrepancy between individual measurements. The intra-correlation coefficient (ICC), assessed by a single grader (absolute agreement, two-way random effect model) was used to reflect the degree of agreement and correlation between measurements. ICCs of <0.50, 0.50-0.75, 0.75-0.90; >0.90 were taken as poor, moderate, good and excellent measures of reliability, respectively [42]. All p-values presented were 2-sided and statistically significant if <0.05.

Quality check

Poor quality scans (low signal strength, presence of motion/blink artefact, improper head positioning etc.) can affect the localization and segmentation performance of the DCNNs, thus resulting in incorrect automated measurements. In this study, we performed a two-step automated quality check based on the predictions obtained to eliminate poor quality ASOCT images. First, upon the detection of the SSL a square region surrounding the center of the predicted region was obtained as the reference. A confidence index was computed as the intersection over union (IoU; between 0-1) between the predicted and reference regions. Scans that yielded a confidence index greater than or equal to 0.85 were considered good, while lower values were designated as poor quality. Second, for the segmentation the number of closed and continuous contours representing each class were used to assess the quality of a scan, i.e., the iris should have two contours, while the corneo-sclera shell and the anterior chamber should have only a single contour each. Scans with predictions that did not satisfy these criteria were considered as poor quality. Finally, the

automatically extracted measurements were considered reliable only if the ASOCT scan satisfied both the aforementioned quality check criteria. The test images are made sure to be of usable quality clinically.

RESULTS

All results in this section are from 4 observers: A – Trained non-expert, B – Trained medical student, C – Fellowship-trained glaucoma expert well-versed in ASOCT analysis, M – Trained machine. The same denotation is used throughout. For the whole study, the mean age ± standard deviation of the patients was 62.20±8.35, the median was 62, the interquartile range was 11 (Q3 = 68, Q1 = 57) and 31.91% of them were males. The percentage for Chinese, Malay, Indian and other races was 77.86%, 11.42%, 7.86% and 2.86% respectively. For testing dataset, the mean age ± standard deviation was 62.00±8.93, the median was 62, the interquartile range was 10 (Q3 = 68, Q1 = 58) and 32.8% of them were males. The percentage for Chinese, Malay, Indian and other races was 75.86%, 15.52%, 8.62% and 0.00% respectively.

Scleral spur localization

First, our proposed segmentation approach was compared against a regression approach, both utilizing DCNNs. The final models were trained for 1,000 iterations and then tested against 3 human observers. The segmentation approach was closer to human observers for all cases. The next test showed that our segmentation approach could reach human level detection with a much smaller training dataset (~200 samples or ~100 images) (**Supplementary Figure 8**).

Inter-observer tests showed that human grader differences were not significantly different from human and machine differences in most cases (**Figure 1A**). Moreover, intraclass correlation [42] (ICC) was done for each observer pair for the X and Y coordinates of the scleral spur location (**Table 2**). It was shown that the machine's scleral spur marking was in high agreement with human graders. Bland-Altman plots for Machine – Human pair was further provided in **Supplementary Figure 9**.

The machine neural network was deterministic once training was complete, meaning that a given input always resulted in the same output. Hence, to do intra-observer tests, another model was trained from scratch and used to compare with the first model. RMS difference for the machine intra-observer test was significantly smaller than most human intra-observer tests (except for observer A, whose intra-observer result was similar to the machine) (**Figure 1B**). This means that machine SSL prediction generally had lower variability than that of human grader.

Table 2. ICC results for Inter Observer Test

Two-way, Single Score, Absolute Agreement ICC									
X Coordinate	A	B	C	M		Y Coordinate	A	B	C	M
A	1	0.978	0.985	0.984		A	1	0.993	0.995	0.994
B		1	0.983	0.979		B		1	0.994	0.993
C			1	0.984		C			1	0.994
M				1		M				1

Figure 1. Observer Test results. A: Inter-observer Test. B: Intra-observer Test

ASOCT segmentation

The ASOCT segmentation performance of the trained network was validated using the Dice coefficient, sensitivity and specificity (**Figure 2**), as described below. The Dice coefficient was used to assess the similarity between the manual segmentation and DCNN segmentation.

The coefficient was defined between 0 and 1 (0: no overlap; 1: perfect overlap), and was calculated for each class as follows:

$$Dice\ score\ = \frac{2\times|D\cap M\,|}{2\times|D\cap M|+|D\setminus M|+|M\setminus D|} \quad [1]$$

where D and M are the set of pixels representing the particular class in the DCNN and manual segmentation, respectively.

Specificity and sensitivity were used to obtain the true negative (assess false predictions) and true positive rates (assess correct predictions) respectively. They were defined for each class as follows:

$$Specificity\ = \frac{|\bar{D}\cap\bar{M}\,|}{|\bar{M}\ |} \quad [2]$$

$$Sensitivity = \frac{|D\cap M\,|}{|M|} \quad [3]$$

Both specificity and sensitivity were defined between 0 and 1.

Figure 2. Validation scores for ASOCT segmentation. Machine segmentation result examples can be found in **Supplementary Figure 10**.

Measurement quantification

Measurement quantification was a crucial step to help validate the scleral spur localization. The segmentation used in this step was fully automated, based on the

assumption that the accuracy of automated ASOCT segmentation is already high. **Figure 3** defined the measured ACA measurements. **Table 3** shows ICC results for inter- and intra-observer test agreement. Inter-observer test results showed good to excellent agreement between observers, especially between machine and human. Moreover, for measurements with relatively lower ICC between machine and human, the human-human counterpart results were similar. Intra-observer test ICC for machine was higher than human, indicating that the machine was more consistent and stable.

Figure 3. ASOCT Measurement Quantification and Definitions. Anterior Chamber Depth (ACD): axial distance between corneal endothelium to anterior lens surface[26]. Lens Vault (LV) : perpendicular distance from the middle of the line connecting the scleral spurs to the anterior pole of the lens[27]. Anterior Chamber Width (ACW): distance between the two scleral spurs[28]. Anterior Chamber Area (ACA): the area bordered by posterior surface of the cornea, anterior surface of iris and anterior surface of the lens[29]. Angle Opening Distance (AOD): distance between the anterior iris surface and posterior corneal surface on a line perpendicular to the trabecular meshwork, at a specific distance from the scleral spur (500µm, 750µm etc.) [30]. Trabecular Iris Space Area (TISA): area of a trapezoid created by the following boundaries: AOD of a distance from scleral spur (500µm, 750µm etc.), line from scleral spur perpendicular to plane of inner scleral wall to the iris, inner corneoscleral wall, iris surface[30]. Iris thickness (IT): IT at a distance from the scleral spur or a relative distance in the iris (e.g.: middle of iris) [31]. Iris Curvature (ICurve): distance from iris greatest convexity point to the line between most central and most peripheral iris pigment epithelium[31].

Table 3. ICC results for Inter and Intra Observer Tests for ASOCT measurement quantification for ACW, TISA and AOD

Inter Observer Test (Two-way, single score, absolute agreement ICC)				
	A vs M	B vs M	C vs M	A vs B vs C
ACW	0.941	0.931	0.949	0.937
TISA500	0.784	0.722	0.710	0.759
TISA750	0.822	0.728	0.761	0.793
AOD500	0.910	0.902	0.927	0.926
AOD750	0.880	0.863	0.898	0.903
Intra Observer Test (Two-way, single score, absolute agreement ICC)				
	M	A	B	C

ACW	0.979	0.951	0.953	0.954
TISA500	0.847	0.845	0.728	0.646
TISA750	0.884	0.887	0.738	0.702
AOD500	0.959	0.958	0.923	0.881
AOD750	0.948	0.956	0.874	0.901

Results visualization and quality check

This was assessed visually by exporting the software prediction into an image format. The machine was able to visualize the per-scan results (**Figure 4A)**. Moreover, fully automated measurement enables 360° analysis, for example of AOD and TISA (**Figures 4B and 4C)**. The goniogram showed that the inferior quadrant's angle is narrower than other quadrants of that specific patient's eye (**Figures 4B and 4C**). Indicating that a global assessment would provide a more accurate diagnosis.

For image quality check, the ASOCT scans need to pass both the SSL confidence and ASOCT segmentation quality assessment. The SSL confidence can be visualized in 360° as shown in **Figure 5A**. Visually comparison of good (**Figure 4A)** and failed (**Figure 5B** and **5C)** cases determined that, if the image quality is good, the SSL confidence should be above 0.85. Detailed analysis to justify confidence threshold to be 0.85 can be found in Appendix A. Moreover, this threshold can be manually adjusted. A failed SSL detection can be seen in **Figure 5B** on the left scleral spur, where SSL confidence is accordingly very low. For ASOCT segmentation, the exclusion criteria are for iris, anterior chamber, corneo-sclera, a number of contours larger than 5, 6 and 10, respectively. Ideally, the number of contours for the said areas of interest should be 2, 1 and 1 respectively. However, for narrow angle cases and many other noisy cases, there might be insignificant wrong small contours. Hence, we increased the threshold. All of these are hyper-parameters and can be tuned. A future systematic study of hyper-parameter tuning is planned. A failed ASOCT segmentation can be seen in **Figure 5C**. All failed scans were excluded from the final measurement quantification.

Figure 4. Example of automated results. (A) Example measurement quantification of a single scan. (B) Example of 360° analysis for AOD. (C) Example 360^{o} analysis for TISA. The measured value for each scan in the whole volume is denoted by the radius, while the angle corresponds to the scan position in the ASOCT volume.

Figure 5. Example of quality check results. (A) Visualization of SSL confidence 360°. Greens are passed scans. Reds are failed scans. Blue circle is 0.8 SSL confidence threshold. Red dots above the thresholds are scans that failed the ASOCT segmentation check. In this example 4/128 scans are disqualified. (B) Excluded scan due to low SSL confidence on the left side. (C) Excluded scan due to bad segmentation quality.

DISCUSSION

The use of ASOCT for the assessment of the ACA in angle closure glaucoma is increasingly popular in the clinical setting. However, the practicality and efficiency of its assessment remains challenging for the ophthalmologists. In the absence of an absolute ground truth for SSL, any prediction, including that of experienced human graders, may be expected to contain errors and show variability in performance. The errors consist of bias, variance and irreducible error (noise) [43 44]. Thus, when a machine learns from human graders, it also learns the human's error. However, with more trainers and data, the errors would be centered around zero [44 45]. In addition, if the algorithm is developed using expert trainers' inputs, these errors would stabilize faster. In clinical practice, errors and variability in SSL on ASOCT scans have huge impact in the diagnosis of angle closure glaucoma because incorrect identification of SSL can result in misdiagnosis and management of patients with PACG. ASOCT imaging has been shown to be more objective and quantifiable compared to gonioscopic techniques [9 13 46 47]. The ACA measurements from ASOCT scans are heavily dependent on the SSL and ophthalmologists gauge treatment effectiveness based on ASOCT measurements before and after treatment.

One of the strengths of the presented method is that it utilizes 3 different approaches to identify the SSL, allowing the machine to be more robust and, thus, be able to more accurately locate the SSL on a variety of ASOCT scans. For ASOCT segmentation, beside a high Dice coefficient, the network also had high sensitivity and specificity, making it a reliable tool in quantifying ASOCT measurements. A comparable algorithm is the STAR Program available on the Casia 2 swept-source ASOCT (Tomey Corporation, Nagoya, Japan), which is capable of automated identification of SSL and ACA measurements [48]. However, this program is a semi-automated software which uses simple edge detection to detect the scleral spur-uvea edge line and, from that, detect the scleral spur location [48]. Moreover, it also depends on the assumption that SSL lies in a perfect circle. In cases of narrow angle, there will be iridotrabecular contact and the scleral spur-uvea edge line will not be visible. In our approach, the machine is trying to learn from human expertise, hence it can detect the scleral spur without the edge line and it also has the potential to expand its definition of scleral spur implicitly by learning from the expert human grader.

The two main limitations in our study were firstly the lack of an absolute ground truth in labelling of the ASOCT images and secondly the size of the dataset. The labelled data was being prepared by human trainers. This is compounded by crowding of the ACA in eyes with

angle closure. The compressed ocular tissues, namely the cornea, peripheral iris and trabecular meshwork, make accurate identification of the scleral spur challenging. Hence, one of the limitations of the paper is the lack of trainers. To validate the machine's performance without a true ground truth, we used the inter- and intra-observer test and ICC, with the exception of the ASOCT segmentation where we only had one trainer and observer. Through the validation tests conducted, it was shown that the machine performance was in good agreement with human performance, while the former was more consistent.

One of the limitations of the study was the relatively small test set, which included a predominantly Chinese population and only one type of ASOCT scan. As such, the generalizability of the results of our study needs to be interpreted with caution outside these circumstances.

As mentioned before, the lack of a generalized population of trainers caused the machine's performance to be biased towards the trainers' errors. As shown in our inter-observer test, since observer A was a trainer for the network, the distance between machine and observer A was lower than the machine with observer B or C. This limitation could be resolved simply by having more trainers. The second limitation was the presence of only one expert. Again, this could be resolve by having more experts.

One technical limitation of our approach was that the resolution depends on the Focus region. The landmarks could not lie too close to the border. The distance should be larger than half of the focus region length, since the point of interest lay in the center of the region. This could be resolved partially with padding (introduce non-meaningful features) or decreasing the size of focus region (susceptible to class imbalances [49]). In this study, the majority of our patients were of Chinese ethnicity (77.86%). It was therefore not possible to perform robust structural comparisons across ethnic groups.

The impact of our method of accurate and automated identification of the scleral spur in ASOCT scans would be in the diagnosis and monitoring of angle closure glaucoma eyes. The diagnosis of angle closure on ASOCT images is dependent on accurate localization of the scleral spur. Angle closure is defined by contact between the peripheral iris and the trabecular meshwork anterior to the scleral spur [9]. As such, the accurate localization of the scleral spur can potentially make screening of angle closure glaucoma on ASOCT imaging easier and more automated. This is especially useful for modern swept-source ASOCT which provides a 360-degree scan of the eye and as many as 64 cross-section cuts of the ACA per eye. The automated identification of the scleral spur reduces variability of human graders and speeds up image analysis to provide a more comprehensive evaluation of the ACA. In the monitoring of angle closure glaucoma eyes, the ACA characteristics should be tracked over time and this paper demonstrates how these measurements can be quantified in a reproducible manner, as most ACA measurements use the scleral spur as the reference. These ACA measurements are important in determining the mechanisms of angle closure, guiding clinical management

and measuring efficacy of treatment modalities [50 51]. In future, the proposed algorithm might make ASOCT scans more clinician-friendly but more studies would be required to determine its diagnostic performance and how it compares to clinical assessments without AI.

Acknowledgements

This work was supported by Singapore Ministry of Education Academic Research Funds Tier 1 (R-397-000-294-114 [MJAG]); Singapore Ministry of Education Tier 2 (R-397-000-280-112, R-397-000-308-112 [MJAG]); National Medical Research Council (NMRC/1442/2016, NMRC/CSA-SI/0012/2017 [C-YC]); National Medical Research Council (NMRC/CNIG/1135/2015 [VK]).

Declaration of Interest

Dr. Michaël J. A. Girard and Dr. Alexandre H. Thiéry are co-founders of Abyss Processing Pte Ltd.

References

1. Tham Y-C, Li X, Wong TY, Quigley HA, Aung T, Cheng C-Y. Global prevalence of glaucoma and projections of glaucoma burden through 2040: a systematic review and meta-analysis. Ophthalmology 2014;**121**(11):2081
2. Quigley HA, Broman AT. The number of people with glaucoma worldwide in 2010 and 2020. The British journal of ophthalmology 2006;**90**(3):262-67 doi: 10.1136/bjo.2005.081224[published Online First: Epub Date]|.
3. Cheng J-W, Zong Y, Zeng Y-Y, Wei R-L. The prevalence of primary angle closure glaucoma in adult Asians: a systematic review and meta-analysis. PloS one 2014;**9**(7):e103222
4. Friedman DS. Epidemiology of angle-closure glaucoma. J Curr Glaucoma Pract 2007;**1**(1):1-3
5. Pan Y, Varma R. Natural history of glaucoma. Indian journal of ophthalmology 2011;**59**(Suppl1):S19
6. See JL, Aquino MCD, Aduan J, Chew PT. Management of angle closure glaucoma. Indian journal of ophthalmology 2011;**59**(Suppl1):S82
7. Casson R, Newland H, Muecke J, et al. Gonioscopy findings and prevalence of occludable angles in a Burmese population: the Meiktila Eye Study. British journal of ophthalmology 2007;**91**(7):856-59
8. Seager FE, Wang J, Arora KS, Quigley HA. The effect of scleral spur identification methods on structural measurements by anterior segment optical coherence tomography. Journal of glaucoma 2014;**23**(1):e29-38 doi: 10.1097/IJG.0b013e31829e55ae[published Online First: Epub Date]|.
9. Nolan WP, See JL, Chew PTK, et al. Detection of Primary Angle Closure Using Anterior Segment Optical Coherence Tomography in Asian Eyes. Ophthalmology 2007;**114**(1):33-39 doi: https://doi.org/10.1016/j.ophtha.2006.05.073[published Online First: Epub Date]|.
10. Sakata LM, Lavanya R, Friedman DS, et al. Comparison of Gonioscopy and Anterior Segment Ocular Coherence Tomography in Detecting Angle Closure in Different Quadrants of the Anterior Chamber Angle. Ophthalmology 2008;**115**(5):769-74 doi: https://doi.org/10.1016/j.ophtha.2007.06.030[published Online First: Epub Date]|.

11. Quek DT, Nongpiur ME, Perera SA, Aung T. Angle imaging: advances and challenges. Indian journal of ophthalmology 2011;**59**(Suppl1):S69
12. Tan AN, Sauren LDC, de Brabander J, et al. Reproducibility of anterior chamber angle measurements with anterior segment optical coherence tomography. Investigative Ophthalmology & Visual Science 2011;**52**(5):2095-99 doi: 10.1167/iovs.10-5872[published Online First: Epub Date]|.
13. Maslin JS, Barkana Y, Dorairaj SK. Anterior segment imaging in glaucoma: an updated review. Indian journal of ophthalmology 2015;**63**(8):630
14. Cumba RJ, Radhakrishnan S, Bell NP, et al. Reproducibility of Scleral Spur Identification and Angle Measurements Using Fourier Domain Anterior Segment Optical Coherence Tomography. Journal of Ophthalmology 2012;**2012**:1-14 doi: 10.1155/2012/487309[published Online First: Epub Date]|.
15. Devalla SK, Chin KS, Mari J-M, et al. A deep learning approach to digitally stain optical coherence tomography images of the optic nerve head. Investigative ophthalmology & visual science 2018;**59**(1):63-74
16. Devalla SK, Renukanand PK, Sreedhar BK, et al. DRUNET: a dilated-residual U-Net deep learning network to segment optic nerve head tissues in optical coherence tomography images. Biomedical optics express 2018;**9**(7):3244-65
17. U-net: Convolutional networks for biomedical image segmentation. International Conference on Medical image computing and computer-assisted intervention; 2015. Springer.
18. Havaei M, Davy A, Warde-Farley D, et al. Brain tumor segmentation with deep neural networks. Medical image analysis 2017;**35**:18-31
19. Li X, Chen H, Qi X, Dou Q, Fu C-W, Heng PA. H-DenseUNet: Hybrid densely connected UNet for liver and liver tumor segmentation from CT volumes. arXiv preprint arXiv:1709.07330 2017
20. Sermanet P, Eigen D, Zhang X, Mathieu M, Fergus R, LeCun Y. Overfeat: Integrated recognition, localization and detection using convolutional networks. arXiv preprint arXiv:1312.6229 2013
21. Deep neural networks for object detection. Advances in neural information processing systems; 2013.
22. Pohlen T, Hermans A, Mathias M, Leibe B. Fullresolution residual networks for semantic segmentation in street scenes. arXiv preprint 2017
23. Zhang J, Liu M, Shen D. Detecting anatomical landmarks from limited medical imaging data using two-stage task-oriented deep neural networks. IEEE Transactions on Image Processing 2017;**26**(10):4753-64
24. Sakata LM, Lavanya R, Friedman DS, et al. Assessment of the scleral spur in anterior segment optical coherence tomography images. Archives of Ophthalmology 2008;**126**(2):181-85
25. Schindelin J, Arganda-Carreras I, Frise E, et al. Fiji: an open-source platform for biological-image analysis. Nature methods 2012;**9**(7):676
26. Eslami Y, Latifi G, Moghimi S, et al. Effect of adjunctive viscogonioplasty on drainage angle status in cataract surgery: a randomized clinical trial. Clinical & experimental ophthalmology 2013;**41**(4):368-78
27. Nongpiur ME, He M, Amerasinghe N, et al. Lens vault, thickness, and position in Chinese subjects with angle closure. Ophthalmology 2011;**118**(3):474-79
28. Nongpiur ME, Sakata LM, Friedman DS, et al. Novel association of smaller anterior chamber width with angle closure in Singaporeans. Ophthalmology 2010;**117**(10):1967-73
29. Wu R-Y, Nongpiur ME, He M-G, et al. Association of narrow angles with anterior chamber area and volume measured with anterior-segment optical coherence tomography. Archives of ophthalmology 2011;**129**(5):569-74
30. Wang D, Qi M, He M, Wu L, Lin S. Ethnic difference of the anterior chamber area and volume and its association with angle width. Investigative ophthalmology & visual science 2012;**53**(6):3139-44

31. Sng CC, Allen JC, Nongpiur ME, et al. Associations of iris structural measurements in a Chinese population: the Singapore Chinese Eye Study. Investigative ophthalmology & visual science 2013;**54**(4):2829-35
32. Han Y, Ye JC. Framing U-Net via deep convolutional framelets: Application to sparse-view CT. IEEE transactions on medical imaging 2018;**37**(6):1418-29
33. 3D U-Net: learning dense volumetric segmentation from sparse annotation. International Conference on Medical Image Computing and Computer-Assisted Intervention; 2016. Springer.
34. Drozdzal M, Vorontsov E, Chartrand G, Kadoury S, Pal C. The importance of skip connections in biomedical image segmentation. Deep Learning and Data Labeling for Medical Applications: Springer, 2016:179-87.
35. Deep residual learning for image recognition. Proceedings of the IEEE conference on computer vision and pattern recognition; 2016.
36. Hansen LK, Salamon P. Neural network ensembles. IEEE Transactions on Pattern Analysis & Machine Intelligence 1990(10):993-1001
37. Marmanis D, Wegner JD, Galliani S, Schindler K, Datcu M, Stilla U. Semantic segmentation of aerial images with an ensemble of CNNs. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences 2016;**3**:473
38. Kingma DP, Ba J. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980 2014
39. Generalized cross entropy loss for training deep neural networks with noisy labels. Advances in Neural Information Processing Systems; 2018.
40. Xu B, Wang N, Chen T, Li M. Empirical evaluation of rectified activations in convolutional network. arXiv preprint arXiv:1505.00853 2015
41. Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R. Dropout: a simple way to prevent neural networks from overfitting. The Journal of Machine Learning Research 2014;**15**(1):1929-58
42. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. Journal of chiropractic medicine 2016;**15**(2):155-63
43. Geman S, Bienenstock E, Doursat R. Neural Networks and the Bias/Variance Dilemma. Neural Computation 1992;**4**(1):1-58 doi: 10.1162/neco.1992.4.1.1[published Online First: Epub Date]|.
44. Shalev-Shwartz S, Ben-David S. *Understanding machine learning: From theory to algorithms*: Cambridge university press, 2014.
45. Barber D. *Bayesian Reasoning and Machine Learning*. Cambridge: Cambridge University Press, 2012.
46. Chansangpetch S, Rojanapongpun P, Lin SC. Anterior segment imaging for angle closure. American journal of ophthalmology 2018;**188**:xvi-xxix
47. Porporato N, Baskaran M, Aung T. Role of anterior segment optical coherence tomography in angle-closure disease: a review. Clinical & experimental ophthalmology 2018;**46**(2):147-57
48. Okamoto K, Higashita R. 2017.
49. Ling CX, Sheng VS. Class imbalance problem. Encyclopedia of machine learning: Springer, 2011:171-71.
50. Moghimi S, Chen R, Hamzeh N, Khatibi N, Lin SC. Qualitative evaluation of anterior segment in angle closure disease using anterior segment optical coherence tomography. Journal of current ophthalmology 2016;**28**(4):170-75
51. Koh V, Keshtkaran MR, Hernstadt D, Aquino MCD, Chew PT, Sng C. Predicting the outcome of laser peripheral iridotomy for primary angle closure suspect eyes using anterior segment optical coherence tomography. Acta ophthalmologica 2019;**97**(1):e57-e63